IOP Conf. Series: Journal of Physics: Conf. Series **1222** (2019) 012045 doi:10.1088/1742-6596/1222/1/012045

On the use of AI based vibration condition monitoring of wind turbine gearboxes

Sofia Koukoura, James Carroll, Alasdair McDonald

University of Strathclyde, Glasgow, UK

E-mail: sofia.koukoura@strath.ac.uk

Abstract. Condition monitoring (CM) systems are installed in wind turbines (WTs) in order to avoid component downtime and reduce maintenance costs. Vibration monitoring is widely used for the WT gearbox, which is a component with a significant downtime. Given that the installed wind capacity grows, the volume of CM data increases, making manual interpretation of vibration signals challenging. Therefore, there is a need for an efficient and automated maintenance decision support system. The aim to this paper is to propose an automated framework for gearbox incipient failure diagnosis. The framework utilises vibration signals and performs health estimation and fault isolation based on signal processing and artificial intelligence (AI) techniques. The methodology is demonstrated through a case study of vibration data from operating WTs with known gearbox failures. The study can be used to optimise wind turbine maintenance actions.

1. Introduction

Operation and maintenance (O&M) costs are a major contributor to the total cost of energy of large wind farms. Wind turbine (WT) failures and unscheduled maintenance actions increase the O&M costs and downtime, compromising therefore the annual energy production. In order to reduce those costs, condition monitoring (CM) systems are installed in wind turbines. The WT gearbox is one of the components with the highest downtime [1]. Vibration based CM has proven to be effective in this component. Accelerometers installed on the gearbox surface transmit high frequency signals that can give useful information about the health state of the gearbox.

A gearbox has various stages composed of gears, bearings and shafts, so the root cause of a gearbox failure can be attributed to many different failure modes. Vibration signals can give some understanding on what causes a failure. Vibration CM in wind turbine gearboxes has been widely researched in the literature. Various existing vibration analysis methods are evaluated and presented in [2]. However, as the installed wind capacity grows, the volume of CM data increases and manual interpretation of wind turbine vibration data becomes challenging. Therefore, in the era of industrial Artificial Intelligence (AI), it is possible to explore if machine learning techniques can aid the decision making process of wind turbine gearbox maintenance. AI has been applied in machine diagnostics [3] and there are a few studies in wind turbines [4], [5].

To this end, the objective of this paper is to present a framework that uses a stream of vibration data from wind turbine gearboxes and automatically isolates incipient faults. The novelty in the paper lies in the automatic wind turbine gearbox diagnostic framework

IOP Conf. Series: Journal of Physics: Conf. Series **1222** (2019) 012045 doi:10.1088/1742-6596/1222/1/012045

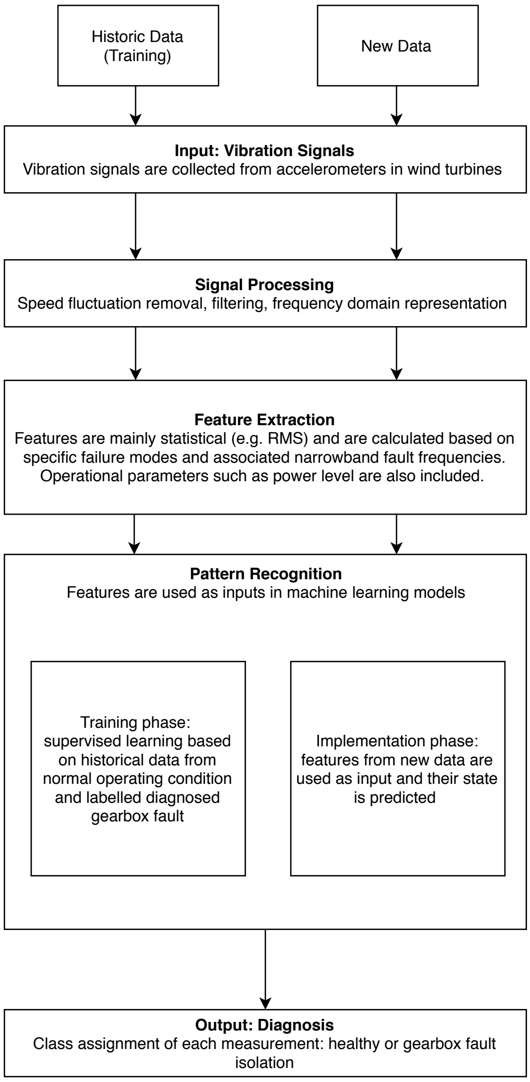


Figure 1. Methodology Flowchart

methodology using domain knowledge and its validation using vibration data from operating wind turbines with confirmed gearbox faults.

2. Methodology

The aim is to create an automated framework that is able to predict and diagnose incipient gearbox failures using AI. The inputs of the models are vibration signals and the outputs are health estimation of the gearbox and component fault isolation and diagnosis. First, signals are collected and processed in order to reveal potential fault signatures. Then, features are extracted from the processed signals around specific frequencies of interest related to gearbox component faults. Finally, these features are used as inputs in pattern recognition models that can determine if a fault is present and in which component and stage it occurs.

An overview of the methodology is given in Figure 1.

IOP Conf. Series: Journal of Physics: Conf. Series **1222** (2019) 012045 doi:10.1088/1742-6596/1222/1/012045

2.1. Signal Processing
Signal processing of the raw time series of vibration signals needs to be applied in order to remove speed fluctuations [6]. Diagnosis through vibration signals can be performed using time, frequency or time-frequency methods [7], with the two latter having shown the most promising results since they make it easier for specific frequency components of interest to be identified. Various techniques can be applied, like Fourier transform, Hilbert transform, cepstrum or envelope spectrum.

Gears faults affect the stiffness of the nearby teeth and this produces changes in the vibration signal. The changes are reflected as amplitude and frequency modulation [8] which can be detected as sidebands around the gear mesh frequencies in the spectrum. A simple sideband energy ratio [9] or cepstrum analysis is often used to diagnose these types of faults. Bearing faults excite high frequency resonances but their signatures are often weaker that the gear ones. A comprehensive tutorial on rolling element bearing diagnostics and advanced signal processing techniques that can be applied is given in [10].

Therefore, since the physics of the faults of the wind turbine gearbox is well understood, any appropriate signal processing method can be used to reveal potential fault signatures.

2.2. Feature Extraction
Features are extracted from historical vibration signals. Instead of using the whole vibration signal, the features provide a more straightforward health indication. A summary of vibration condition indicators can be found in [11]. These features are mainly statistical (e.g. RMS) and are calculated based on specific failure modes and associated narrowband fault frequencies. The operating conditions have a strong effect on vibration signals, so variables such as the power and operating speed are also considered as features.

2.3. Pattern Recognition
The aforementioned features are used as inputs in pattern recognition models. There exists a wide variety of supervised machine learning models that the interested reader can choose from [12]. These models are trained based on historical data, from both normal and abnormal operation, with the aim of predicting gearbox faults. The model is essentially a multi class classifier. The classes are the diagnostic state of the gearbox; either "healthy" or the failure mode of a gearbox component. The number of classes therefore depends on the number of failure mode examples used to train the machine learning model.

When new data samples become available, they are processed in a similar way as the historic signals used for training and the features extracted are used as inputs in the trained pattern recognition model. The output of the model is the classification of the sample that informs whether the gearbox is healthy or where a potential failure is located.

3. Case Study
Three wind turbines with three confirmed gearbox failures are used for the present case study. The wind turbines are offshore, same model and are rated between 2 and 4 MW. The gearbox examined is a double planetary stage with one parallel stage. The three confirmed gearbox faults examined are:

(i) High speed gear crack
(ii) Intermediate speed gear crack
(iii) High speed bearing outer race fault

A schematic is shown in Figure 2 with fault locations and the approximate accelerometer and tachometer locations.

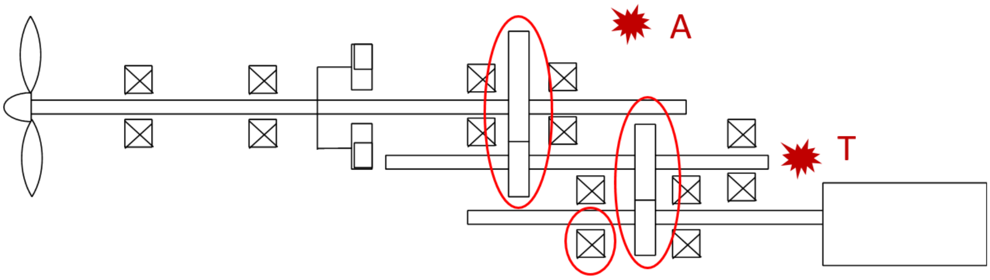
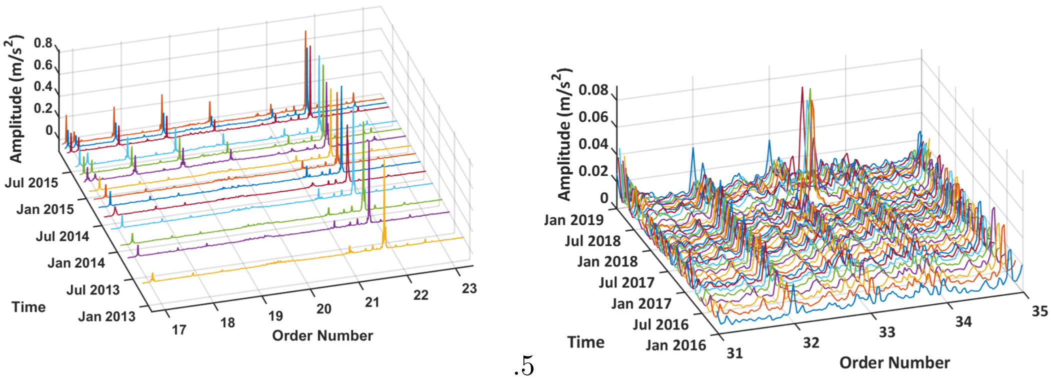


Figure 2. Gearbox diagram and fault locations.

Vibration signals are collected from an accelerometer positioned at the top of the gearbox close to the high speed stage. A tachometer pulse installed on the high speed shaft is used to resample the signals and they are transformed in the order domain. The results from processed signals progressively before failure are shown in Figure 3.



.5 .5

Figure 3. Order spectra leading up to a high speed pinion fault (right) and high speed bearing fault (left). Narrowband failure frequencies are shown, for gear mesh frequency and 5th harmonic of ball passing frequency

The features extracted are shown in Table 1. The gear meshing frequencies and the bearing fault frequencies are known parameters, as long as the design of the gearbox is known. The gear faults are located on the high speed and intermediate speed gears, so the gear frequencies considered are the high speed and intermediate speed gear meshing frequencies. Regarding the high speed bearing, the 5th harmonic of the ball passing frequency outer is used, because through observation of the signals it was concluded by the authors that it gives a good fault indication.

The sideband energy ratio sums the amplitudes of the first six sideband peaks on each side of the gear mesh frequency $\sum_{i=-6}^{6} A_{SB,i}$ and divides by the amplitude of the center mesh frequency A_F. For a healthy gear mesh the sidebands have a small amplitude compared to the center mesh frequency. As damage develops on a gear tooth, the sidebands rise in amplitude which results in a larger sideband energy ratio value. In a similar way, the energy of the bearing ball passing frequency is calculated, by considering a vector A which is centered at the bearing fault frequency and a window +-5% around it. The value is normalised with the length of the data segment. The RMS value can also provide a similar indicator and is calculated in a vector centered the gear mesh frequencies or bearing fault frequencies of interest and a window +-5% around it.

Table 1. Calculated features from the order domain and operational parameters used as inputs in the machine learning model.

	Feature	**Calculation**
Gear fault	RMS	$\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}A(i)^2}$
	Sideband Energy Ratio	$\frac{\sum_{i=-6}^{6}A_{SB,i}}{A_F}$
Bearing fault	RMS	$\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}A(i)^2}$
	Crest Factor	$\frac{max\lvert A(i)\rvert}{\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}A(i)^2}}$
Operational	Power	Provided channel
	Speed	Provided channel

Table 2. Confusion matrix of gearbox fault classification.

Actual \ Predicted	**HS Bearing Fault**	**IS Gear Fault**	**HS Gear Fault**	**Healthy**
HS Bearing Fault	92%	<1%	<1%	7%
IS Gear Fault	1%	84%	4%	11%
HS Gear Fault	<1%	<1%	94%	5%
Healthy	2%	7%	4%	87%

Bootstrap-aggregated decision trees are chosen because they have proven to be a robust classifier [13]. They combine the results of many decision trees, thus reducing the effects of overfitting and improving generalization. The data is split into training and testing with a ratio 80/20%. The training set is used to build the classifier. The hyperparameters tuned are the maximum number of trees, the depth of the trees and the maximum number of features considered for splitting a node. These hyperparameters are tuned using random grid search and cross validation. Supervised learning is performed using labeled data. The labels are the health state of the component or fault isolation. Each sample is labeled as "healthy" if there was no incipient fault signature present in the vibration signals. If there are any frequencies present, indicating the developing of a fault, the signal is classified according to the type of failure mode, as enlisted above. The dataset was fairly balanced after downsampling the "healthy" data points.

The results of the multi class classifier for the testing dataset are shown in the confusion matrix Table 2. The vertical axis represents the actual values and the horizontal one represents the predicted ones. The aim is to have as many accurate predictions as possible (lots of points in the diagonal). Each row shows how the data of the testing set for each label are predicted using the machine learning model trained on the training set.

According to these results, there is a fairly accurate classification of the health state of the gearbox. It should be noted that the accuracy of the machine learning algorithm often depends on the quality of the features, and in this case the domain knowledge of the physics of failure of faults is used to extract features that are good condition indicators.

4. Conclusion

This paper presented an automated methodology for processing vibration signals, extracting features and using machine learning models to isolate incipient gearbox faults. The promising results suggest that AI can be used in order to make more informed maintenance decisions and avoid manual interpretation of a large amount of wind turbine CM data. The methodology is valid for vibration CM of WT gearboxes and could be utilised for optimising O&M actions. It should be noted however that sufficient number of training data and failure examples should be used to create pattern recognition models. Future research will focus on case studies of more failure modes and comparison of the suggested methodology against existing techniques used in industry.

5. Acknowledgements

The authors would like to acknowledge EPSRC funding number EP/L016680/1 for the funding of this project.

6. References

[1] James Carroll, Alasdair McDonald, and David McMillan. Failure rate, repair time and unscheduled o&m cost analysis of offshore wind turbines. *Wind Energy*, 19(6):1107–1119, 2016.
[2] Shuangwen Sheng. Wind turbine gearbox condition monitoring round robin study-vibration analysis. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.
[3] Arfat Siddique, GS Yadava, and Bhim Singh. Applications of artificial intelligence techniques for induction machine stator fault diagnostics. 2003.
[4] Pedro Santos, Luisa F Villa, Aníbal Reñones, Andres Bustillo, and Jesús Maudes. An svm-based solution for fault detection in wind turbines. *Sensors*, 15(3):5627–5648, 2015.
[5] Sofia Koukoura, James Carroll, and Alasdair McDonald. Wind turbine intelligent gear fault identification. In *Annual Conference of the PHM Society*, 2017.
[6] KR Fyfe and EDS Munck. Analysis of computed order tracking. *Mechanical Systems and Signal Processing*, 11(2):187–205, 1997.
[7] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510, 2006.
[8] PD McFadden. Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *Journal of vibration, acoustics, stress, and reliability in design*, 108(2):165–170, 1986.
[9] J Hanna, C Hatch, M Kalb, A Weiss, and H Luo. Detection of wind turbine gear tooth defects using sideband energy ratio^{TM}. *China Wind Power 2011; October, 19-21, 2011, Beijing, China*, 2011.
[10] Robert B Randall and Jerome Antoni. Rolling element bearing diagnosticsâa tutorial. *Mechanical systems and signal processing*, 25(2):485–520, 2011.
[11] Junda Zhu, Tom Nostrand, Cody Spiegel, and Brogan Morton. Survey of condition indicators for condition monitoring systems. In *Annu. Conf. Progn. Heal. Manag. Soc*, volume 5, pages 1–13, 2014.
[12] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, NY, USA:, 2001.
[13] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.